

be improved.

Further, generally, there is the possibility that the insertion of the complementary packet PP will lower the quality of the decoded voice output. Therefore, it is desirable not to insert the complementary packet PP, if possible. As described in the problem that the invention intends to solve, the packet is inserted only when the queue length becomes zero in the case where it is inserted only into the top position, and therefore it can be said that this is a method for reducing the insertion frequency of the complementary packet PP as much as possible.

On the other hand, in this embodiment, the complementary packet PP is inserted when the queue length becomes less than the lower threshold TL1 that is not zero, and therefore the possibility that the insertion frequency will rise is high, and, according to circumstances, there is a possibility of inserting the complementary packet PP that should not be inserted.

However, in this embodiment, since the complementary packet PP is inserted so as to be dispersed while selecting the insertion position thereof even if the insertion frequency thereof is high, a deterioration in the quality of the decoded voice output caused by the insertion is very slight if any, and it is possible to prevent a marked deterioration in quality that is caused when the complementary packets PP are continuously inserted, for example.

Preferably, in this embodiment, the lower threshold TL1 is set at as low a position as possible (i.e., near the top position), in order not to insert the complementary packet PP as much as possible. However, it is advantageous to set the lower threshold TL1 at as high a position as possible, in order to insert the packet to be dispersed while selecting an optimum insertion position.

If the queue length continues to be in a state of being below the lower threshold TL1 over a plurality of periods of decoding unit time, a plurality of complementary packets PP must be inserted for a queue whose length is less than 7 packets (at this time, the queue length is shortened by one packet for a lapse of one period of decoding unit time) even if one complementary packet PP is inserted during one period of decoding unit time. Therefore, in this embodiment, there is the possibility that, after all, the complementary packets PP are continuously inserted or necessarily inserted into the interval of voice presence.

However, according to this embodiment, it is possible to reduce the occurrence frequency of the successive insertions or the insertion into the interval of voice presence.

Generally, in a device for decoding a voice by the packet unit, a packet must be decoded at intervals of a fixed decoding unit time. However, according to this embodiment, a voice presence/absence judgment for a packet and the insertion of a packet can precede other processing.

That is, since they can be processed when the load of a processor is low, the operating ratio of the processor can be improved.

#### (E-2) Effect of the fifth embodiment

According to this embodiment, the same effect as that of the first embodiment can be obtained.

In addition, according to this embodiment, it is possible to reduce the occurrence frequency of the continuous insertion of complementary packets (PP) or the insertion into the interval of voice presence, and therefore the possibility of considerably lowering the quality of the decoded voice output can be reduced.

#### (F) Other Embodiments

The characterizing parts of the first through fifth embodiments can be variously combined.

For example, the lower threshold TL1 that is set at a higher position than the top position, which is a feature of the fifth embodiment, can be combined with not only the first embodiment but also the second through fourth embodiments.

Further, although the lower threshold TL1 is fixed in the fifth embodiment, the lower threshold TL1 may be dynamically changed like the higher threshold TH1 of the third embodiment or like the higher threshold TH2 of the fourth embodiment.

In the first through fifth embodiments, data that corresponds to a conversational voice is contained in a